Towards the Development of a Metadata-driven Usability Awareness Prototype for Interoperable GIS Operation Design

Jung-Hong Hong[1], How-Han Chang[1], Ting-Yi Lee[1], Ting-Yu Chang[1]
[1] Dept. of Geomatics, National Cheng Kung University, Tainan, Taiwan,
junghong@mail.ncku.edu.tw ,{P66134072; P66144027;P66134098}@gs.ncku.edu.tw

Keywords: Usability, Awareness, Metadata, Data Quality, GIS Operations

Abstract

The effective functioning of Geographic Information Systems (GIS) relies on three essential components: right data, right operations, and right results. Although challenges associated with data discovery and acquisition have lessened over time, concerns regarding data fitness for intended use and usability assessment remain prevalent. Traditionally, reliance was placed on users’ expertise in understanding the data, executing operations, and professionally interpreting outcomes. However, this conventional methodology is inadequate for managing the extensive range of unfamiliar resources currently available. From a usability perspective, this study introduces an intelligent framework based on standardized metadata design, formalized knowledge from data producers, and automated awareness mechanisms embedded within software systems. The framework aims to effectively facilitate interaction between resource providers and users concerning data usability, thereby improving the resource-sharing environment. It utilizes international standards ISO 19115 and ISO 19157 to select metadata elements relevant to usability, encompassing multiple dimensions such as the universe of discourse, qualitative and quantitative quality, and usability scenarios. This standardized metadata is rigorously analyzed in the context of the intended purposes of GIS data and operation, with a focus on incorporating usability interpretation into the design of operations. When dealing with heterogeneous data sources, this approach supports users in developing accurate comprehension and avoiding misuse, ultimately promoting interoperability and enhancing reusability.

1. Introduction

In recent years, the development trajectory of geographic information system (GIS) technology—across both governmental and private sectors—has increasingly emphasized the flexible integration and utilization of diverse information to deliver high-quality services. Such advancements necessitate adherence to the principle of “the right data, the right operations, and the right results.” From a governmental standpoint, innovative paradigms introduced in recent years—such as digital transformation, evidence-based governance, and smart nation initiatives—underscore a strong commitment to leveraging information technology to optimize resource management models at various administrative levels and to enhance decision-making quality. Concurrently, the government promotes and develops a range of human-centered, sustainable environment, and cross-domain integrated application services aimed at improving citizens’ quality of life and facilitating more effective national land use planning.

A significant advantage of the geographic information application environment lies in its capacity to enable users to access and utilize data shared across different organizational units and domains, thereby fostering enhanced integration benefits. The Technical Committee 211 of International Organization for Standardization (ISO/TC211) and the Open Geospatial Consortium (OGC) have advocated the OpenGIS strategy, which promotes the sharing of geographic data within a standardized, open environment. While individual units may maintain their internal operational modes, data exchange must conform to standardized technical specifications to support nationwide or regional data sharing. This approach aims to surmount various barriers to data sharing and to improve interoperability challenges encountered by users when handling heterogeneous datasets (Manso-Callejo et al., 2009).

Prior to the widespread development of internet technologies, data sharing was already acknowledged as a significant challenge. Over the last twenty years, advancements in internet technology have facilitated the emergence of numerous innovative concepts. Among these, the FAIR principles, introduced by Wilkinson et al. (2016), have attracted substantial scholarly attention. Initially articulated as fifteen guidelines for the sharing of scientific data, these principles emphasize four fundamental dimensions: Findability, Accessibility, Interoperability, and Reusability. They promote data sharing and utilization under conditions that are machine-actionable, underscoring the essential role of data in contemporary research.

Considerable efforts have been directed toward implementing the FAIR principles within Geographic Information Systems (GIS) communities. For instance, Giuliani et al. (2021) developed the SDI-based SwissEnvEO system to facilitate the sharing of large-scale Earth Observation data. By leveraging an extensive data repository, this system integrates FAIR principles into its sharing mechanisms through various geographic web services and metadata management software that adhere to international standards, such as GeoNetwork, thereby enabling the publication of ready-for-use shared resources. The necessity to manage large-scale datasets has also led to the adoption of frameworks such as Data Cube and Digital Twins in specialized domains (Barzaghi et al., 2024). These initiatives aim to realize the FAIR objectives through innovative technological approaches, including the application of open linked data to satisfy accessibility requirements.

Ivánová et al. (2019) contend that, despite the existence of standards such as ISO and OGC, their original design frameworks do not fully align with the requirements of the FAIR principles, indicating the need for further concrete modifications or adaptation strategies. Metadata—commonly defined as “data about data” or “data about georesources”—has long been recognized as a critical component in data sharing frameworks across diverse disciplines (Aalders, 2005). Metadata provides detailed descriptions of various data attributes, thereby facilitating large-scale, cross-domain information management in accordance with FAIR principles.

Within the geographic information domain, ISO 19115 (ISO, 2014) and ISO 19157 (ISO, 2023b) are two widely adopted series of standards, providing common technical references for standardized descriptions of metadata and quality of geographic information (Brodeur et al., 2019). The prevalent strategy involves developing metadata-driven support frameworks for geographic information therefore grounded in ISO 19115 and ISO 19157. For example, Fischer et al. (2023) investigated the integration of standardized quality information into user-driven mechanisms to facilitate suitability assessments. A characteristic feature of these approaches is the initial construction of standardized metadata, followed by content extraction or mapping tailored to specific application scenarios, thereby enabling diverse operations within that environment.

ISO 19157 offers comprehensive rules for resource sharing models and standardized quality evaluation and description. Two principal participants, producers/providers of resources and users demanding resources, corresponding to “supply” and “demand” roles, are respectively identified. To ensure professional data production and avoid redundant investments, establishing effective communication channels between these parties is essential. However, their domain data are often designed according to internal requirements, typically characterized by the “universe of discourse”—the idealized data status defined by responsible units concerning data specifications and content. This encompasses definitions of data objects, filtering criteria (e.g., type, size, width), spatial models, and quality threshold requirements, all of which can be articulated through data product specifications. When content and requirements are concretely defined via such specifications and data are created accordingly, the standardized specification not only provides a basis for data quality verification and description but also offers critical references for user units in interpreting the data.

The assessment of dataset fitness for use has long been recognized as a significant issue (Ahonen-Rainio & Kraak, 2005; Devillers et al., 2007b), with early research predominantly focusing on evaluations grounded in data quality (Devillers et al., 2010). The concept of Quality-Aware GIS is developed upon GIS applications incorporating standardized quality information. Devillers and Zarger (2009) combined standardized quality information with GIS functionalities, proposing mechanisms for reporting and assessing data quality. Hong and Hwang (2017) proposed mechanisms to automatically determine valid spatial extents within GIS interfaces based on data completeness descriptions. Such awareness design can alert users to operational risks, functioning analogously to expert guidance, thereby enhancing the accuracy of user operations.

Although geographic data users are expected to possess the capacity to evaluate whether datasets satisfy specified purposes, unfamiliarity with data content and the lack of professional expertise often impede accurate understanding, potentially resulting in latent and unrecognized risks in subsequent decision-making processes (Devillers et al., 2007a). Despite the widespread adoption of geographic information technologies, determination regarding data fitness for use and correct usage remain heavily reliant on metadata provided by data producers. However, if data producers themselves lack such requisite capabilities, achieving the objectives of the FAIR principles becomes nearly unattainable. From the perspective of usability, this study emphasizes the importance of facilitating the effective transmission of professional knowledge and ensuring correct interoperable application on the user side. The subsequent sections will elaborate on the framework’s design strategies, research findings, and conclusions.

2. Metadata-driven Framework Design

2.1 Design strategies

This research asserts that an optimal environment for resource sharing, which integrates usability considerations, must address the following critical components:

1. Data providers should conduct thorough evaluations of their data products and utilize standardized metadata elements to explicitly convey the characteristics and usage constraints of these products.
2. Metadata produced by interdisciplinary organizations ought to be developed in alignment with international standards and incorporated into a national-level resource sharing system. The adoption of such a standardized framework facilitates the effective implementation of the FAIR principles, thereby improving the efficiency and effectiveness of cross-disciplinary professional resource sharing.
3. User software should embed awareness mechanisms capable of automatically interpreting standardized metadata to enhance users’ comprehension of the data. Additionally, usage restrictions should be communicated through suitable visualization methods to minimize the risk of improper data use or manipulation.

Building upon this ideal framework, the study delineates three principal implementation phases and strategies, concentrating respectively on the analysis and design of standardized metadata elements, the formulation of metadata content compliant with both international and domestic standards, and the integration of these elements within user-end application environments (Figure 1).

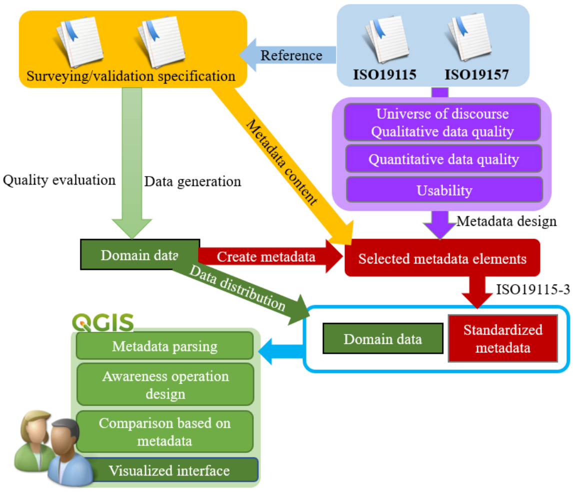

Figure 1 The workflow of the proposed prototype.

2.2 Standardized Metadata Design for Usability

In addressing the need to establish standardized metadata content, this study employs metadata elements sourced from the established schemas of two international standard series, ISO

19115 and ISO 19157. After identifying and synthesizing metadata elements relevant to the characterization of "usability", the research advances to propose recommended guidelines for these descriptions. The choice of metadata elements is justified by their recognition as internationally accepted standards, facilitating the creation of a consensus-based standardized metadata framework and enhancing the potential for intelligent filtering.

Standardized metadata descriptions serve as an interface between providers and users, facilitating guidance and evaluation regarding the suitability of data for specific applications. The considerations are categorized into three primary information domains:

1. Universe of discourse and qualitative data quality
Domain data explicitly delineates the subject matter, constraints, and measurement methodologies employed during data creation, all of which critically influence data selection and usability. This domain content is preferred to reference established data product specifications, with particular emphasis on data object definitions, constraints, and surveying procedures, which are subsequently integrated into the proposed metadata elements. Qualitative quality descriptors predominantly encompass Lineage information, comprising Process Steps and Source references. The integration of these components enables the construction of a comprehensive workflow narrative aligned with the sequential execution of processes.

2. Quantitative Quality
Quantitative quality aims to numerically characterize geographic data quality from multiple perspectives. This study primarily adopts the quality framework outlined in ISO 19157, which encompasses six categories: completeness, logical consistency, positional accuracy, thematic accuracy, temporal accuracy, and usability. ISO 19157 offers standardized quantitative quality descriptors applicable across all the six quality concepts. This standardization facilitates the development of data processing modules for end-users and supports data filtering based on threshold criteria, thereby aiding the selection of right data.

3. Usability
Usability is directly conveyed through the specification of applicable and non-applicable scenarios, employing whitelist and blacklist approaches. Typically, data products are designed with explicit intended uses and corresponding applicable scenarios, which are incorporated into the whitelist. Conversely, scenarios where data usage is inappropriate—due to factors such as purpose, legal constraints, domain specifications, or quality limitations—are identified via the blacklist.

By incorporating the influences of domain characteristics and quality assessments on data usability, the metadata framework proposed herein not only provides clear guidance for data selection but also enhances users’ professional comprehension of the operational conditions and theoretical considerations requisite for effective data utilization.

2.3 Recording of Standardized Metadata

To improve the dissemination and integrated utilization of interdisciplinary information, geographic resources exchanged among various entities should be encoded and distributed utilizing OpenGIS technologies. Recent developments have demonstrated that data formats such as GML, KML, and GeoJSON effectively satisfy these operational requirements. Beyond the framework and semantic information, which must conform to data product specifications or established data standards, the selection of metadata elements—including usability factors examined in this study—should also be documented and shared through open data formats. This study employs the GeoNetwork software, capitalizing on its functionality to create and export ISO 19115-3 metadata (ISO, 2023a), thereby addressing the need for standardized and interoperable data interpretation.

2.4 Awareness GIS Operation Design

Upon acquiring standardized metadata, the user-end software is designed to automatically parse and interpret this data, thereby offering users review and reference support during related operations. The developed prototype aims to assimilate usability information into the automatic interpretation process within GIS software by introducing novel feature developments. The design and implementation of these related operations will be realized as an extension within the QGIS platform. Testing and analysis will incorporate the metadata recommended by this study, embedding them into various operations to achieve automatic and awareness functionalities. The proposed development model encompasses three fundamental strategies:

1. Automatic Parsing and Display of Usability Information
This approach involves the automated extraction and display of usability information derived from metadata when users activate particular functions. For example, it is designed to present critical domain-specific information and delineate suitable or unsuitable application contexts. This method usually provides preliminary support to users during the data selection process.

2. Functional Design Incorporating Usability Information
This methodology entails the restructuring of operational processes that may be affected by usability considerations through the incorporation of usability data interpretation and impact evaluation. It modifies workflow procedures and offers alerts or recommendations from a quality assurance standpoint concerning the results of these operations.

3. Comparative Usability Information for Multiple Datasets
Beyond presenting usability information for individual datasets, this strategy introduces a tabular comparison mode when multiple datasets are concurrently accessed. This mode displays interpretation data results for multiple datasets in a comparative table format, facilitating evaluation based on specific criteria.

The proposed user interface serves as the medium through which users access and select datasets within GIS software. In the present study, a pop-up window-based method was utilized to deliver clearer and more informative feedback. This method proved to be as effective as the ongoing expert guidance typically provided during data selection, thereby alleviating challenges related to the interpretation of large and unfamiliar datasets.

3. Result analysis

3.1 Proposed metadata elements for usability

Cross-disciplinary discovery of geographic data typically depend on constraints based on time, location, theme and responsible authorities. The following sections elaborate on the rationale and supporting explanations for these key metadata categories:

1. Temporal Extent:
Temporal information is documented within the standardized Extent metadata category, which delineates the temporal coverage of the dataset. Inquiries pertaining to temporal data may also require consideration of the semantic context of the associated time-related attributes.

2. Spatial Extent:
Spatial extent is also documented within the Extent metadata class. The spatial extent may be characterized using textual descriptions, bounding coordinates (bounding boxes), and polygonal representations, with coordinate systems standardized to the WGS84 geospatial reference system. While the bounding rectangle is commonly employed, it provides only an approximate representation of spatial extent and may lead to misinterpretations regarding usability.

3. Thematic consideration:
ISO 19115 delineates two components—TopicCategory and Keyword—for the characterization of data themes. The TopicCategory element utilizes a standardized codelist comprising broad thematic domains, functioning chiefly as a preliminary filtering tool given its relatively coarse thematic granularity. In contrast, Keywords afford enhanced flexibility by enabling data producers to attribute descriptive terms that encapsulate more detailed thematic and geographic attributes. The choice of keywords significantly impacts subsequent retrieval processes and may present semantic ambiguities. To facilitate the advancement of national spatial data infrastructures, it is advisable that each thematic domain develop distinct, consensus-driven vocabularies and disseminate them as authoritative references for prospective users.

4. Responsible Authority:
Each data category must be assigned a designated authority accountable for the design, creation, maintenance, and provision of services related to the data. Within the ISO 19115 standard, the Citation category encompasses a range of reference and citation information. Correspondingly, the CI_Responsibility element documents details including the name, contact individual, telephone number, and email address of the responsible organization. Additionally, metadata elements such as credit and PointOfContact serve to identify contributors or relevant contact entities.

However, the aforementioned four descriptive categories alone are insufficient for enabling users to make well-informed assessments concerning the appropriateness of data. To address this limitation, the inclusion of additional metadata elements is intended to offer thorough support for filtering decisions related to data utilization, while simultaneously accounting for potential risks. The selection process is grounded in standardized metadata elements derived from ISO 19115(ISO, 2014) and ISO 19157(ISO, 2023b), with content contributions from service providers who apply their professional expertise regarding resource characteristics and quality. Three primary perspectives guiding the selection of usability metadata are proposed, and Table 1 enumerates the recommended metadata elements.

1. Data Characteristics Perspective:
(1) Data definition: A clear textual explanation of the data theme’s meaning, aiding users in understanding the described phenomena. Ideally, this should reference domain specification documents and incorporate domain keywords to support the development of domain knowledge ontologies.
(2) Data Model: The data model defines the spatial representation of geographic information. Various spatial data models are appropriate for different contexts and should be thoroughly described, incorporating chosen parameters where applicable.
(3) Spatial modelling unit: The spatial modelling unit forms the basis for feature object creation, carries specific semantic meaning, and affects related attributes, applicable operations, and potential outcomes.
(4) Data Version: For datasets subject to ongoing updates, each version should be documented with update timestamps and references to facilitate version selection.

2. Technical Specification Perspective:
(1) Specification Document: If data production and quality assurance follow specific technical specifications, datasets adhering to the same specifications will exhibit consistent content, procedures, and quality, simplifying suitability assessments.
(2) Original Source Data: The source data used during dataset creation critically influences the final product, especially when similar datasets that can be derived from different sources with varying quality or specifications (e.g., Level of Detail 1 building data in this study).
(3) Filtering constraints for creating data: Criteria applied during data creation will impact the final results, such as thresholds on the size, width, or length of the objects, which determine whether certain features are included or excluded.
(4) Data production procedures: Procedural details governing data creation also has impacts. This may specify original data sources and operational steps, including georeferencing, generalization rules, and attribute surveys.
(5) Data product Quality: Despite adherence to production regulations, data quality must be verified prior to release. This information is suitable for description using quantitative quality metrics.

3. Analysis of Suitable and Unsuitable Applications:
(1) Original Intended Application: Explicitly states the scenarios for which the data was originally designed and created.
(2) Extendable Application Scenarios: Describes additional applications beyond the original intent where the data may be suitably applied, often based on successful use cases or quality criteria. This represents a whitelist approach.
(3) Unsuitable Application Scenarios: Identifies limitations arising from data creation considerations, highlighting scenarios where the data is inappropriate or requires caution. This represents a blacklist approach and provides professional guidance on data applicability.

Table 1 Recommended usability metadata elements.

Element	Content recommendation
Contact	Contact information of the organization or individual responsible for building the metadata.
Citation	The name of the dataset, which can include information such as its name, time, spatial coverage, specifications, and issuing organization.
Abstract	A summary description of the dataset, provided in free text, which may

Element	Content recommendation
	include important characteristics of the described dataset.
pointOfContact	The contact organization or individual for the dataset; multiple contacts can be recorded with different roles.
spatialRepresentationType	The basic type of spatial representation, which may include vector, grid, etc., each with different specifications in spatial representation.
spatialResolution	The spatial resolution of the dataset, including scale, planar, elevation and angular spacing, LOD, and other descriptive parameters.
topicCategory	ISO19115 recommended global geographic data theme categories, provided in codelist format.
additionalDocumentation	Reference documents related to the dataset besides the original dataset reference standards.
Keyword	Keywords used to describe different perspectives of the dataset.
Type	Description of the type of keyword selected.
thesauraiName	The source reference of the keywords, often provided as domain-specific reference documents.
specificUsage	Description of the actual application of the dataset in specific scenarios.
userDeterminedLimitations	Situations where the dataset is deemed unsuitable for use by the user.
additionalDocumentation	Supporting documents related to the assessment of data usability.
uselimitation	Restrictions on the use of the dataset, describing limitations under general conditions.
useConstraints	Use restrictions described via codelist, such as protection under copyright, intellectual property rights, and other regulations.
classification	Confidentiality level of the data described via codelist, distinguishing different levels that affect the usage scenarios of the data.
statement	A free-text description of the data processing history of processing the dataset.
additionalDocumentation	Supporting information about referenced standards.
source	Specifies the data used during the data processing history.
processStep	Describes each step in the data processing history; multiple steps can be included, each specifying the data being used.
featureType	Describes the type and basic information of the feature category design.
featureCatelogue	Describes the documentation

Element	Content recommendation
Citation	established for the feature category and catalogue.
EX_geographicExtent	Describe the spatial extent of the dataset, including polygons, bounding coordinates, and place name descriptions.
EX_temporalExtent	Used to represent the temporal extent of the dataset, indicating the state of geographic data content within the valid time period.
EX_verticalExtent	Used to represent the elevation extent of the dataset, applicable to the elevation description of 3D data.
Data completeness	Describes the completeness status of the dataset, which can be set as a filtering threshold.
Logical consistency	Describes the conditions met by the dataset, which can be set as mandatory checks.
Positional accuracy	Describes the positional accuracy of the dataset, which can be set as a filtering threshold.
Thematic accuracy	Describes the accuracy of specific attributes, which can be set as a filtering threshold.
Time quality	Describes the accuracy of temporal descriptions, which can be set as a filtering threshold.
Usability	Describes the appropriate or inappropriate usage scenarios of the dataset, providing content to guide decision-making.

In summary, the selection of these metadata elements enhances users’ ability to assess data suitability comprehensively, thereby supporting informed decision-making in geographic data utilization.

3.2 Example: Usability information of LOD1 building

This section employs Level of Detail 1 (LOD1) buildings as a case study to examine the establishment of usability metadata and its influence on the design of operational processes. The specifications and guidelines for LOD1 buildings are often aligned with the criteria set forth in the CityGML standard. Typically, users engage with the visualized representations of these models without possessing a thorough understanding of the underlying data generation processes or detailed knowledge of their inherent attributes. Variability in LOD1 building models can arise from multiple factors, including the requirement for explicitly defined data themes, differing specifications due to heterogeneous data origins, variations in building selection criteria, the choice of spatial modeling units (such as individual buildings versus townhouses), positional accuracy, diverse methodologies for determining elevation data, temporal discrepancies, surveying techniques, and the granularity of semantic objects, among other factors. Notably, even when datasets are uniformly categorized as LOD1 buildings, these factors can introduce significant heterogeneity, thereby affecting evaluations of data usability.

The following discussion examines, from a usability perspective, the essential understanding users must acquire regarding LOD1 buildings, with the aim of facilitating system-assisted decision-making during data selection and practical application:

1. Spatial Visualization Outcomes
While visualization in GIS interface offers an intuitive means of interaction, it does not inherently reveal possible differences in the specifications or quality of LOD1 building data. Consequently, any operation involving spatial display must account for these considerations.

2. Temporal Update Considerations
Three-dimensional building datasets are generally not maintained with real-time updates. Consequently, any temporal operations require a comprehensive understanding of the temporal semantics linked to time-related attributes, including the data creation date, building completion date, and building age. These temporal characteristics are critical for all operations involving historical versioning and semantic temporal analysis.

3. Variability in LOD1 Building Specifications
Differences in the source domains of building data influence the resulting models and may introduce regional disparities. Even within the LOD1 classification, the level of descriptive detail can vary significantly—for example, in the identification of minor roof structures and roof types—affecting all operations related to the spatial representation of buildings.

4. Feature Count Analysis
The definitions and domain criteria applied to the establishment of data impact the calculation of the number of building features that satisfy specific constraints.

5. Feature Volume and Area Metrics
LOD1 buildings are typically represented as block models. Consequently, their area and volume measurements may not conform to legal or application requirements, influencing all operations that utilize these spatial metrics.

6. Cross-Domain Spatial Data Analysis
When LOD1 buildings are developed independently, they may lack accurate positional or temporal relationships with other thematic datasets (e.g., cadastral information). Evaluations of from both positional and temporal aspects must be conducted jointly, as these factors affect all spatial analyses involving multiple data types and the coherence between datasets.

7. Height and Floor Number Assessments
The height information in LOD1 building data may derive from various sources, including direct measurements, Digital Surface Models (DSMs), or calculations based on registered floor counts, each with differing accuracy levels. These discrepancies influence all height-dependent analyses, such as viewshed or shadow assessments. Furthermore, discrepancies can occur when integrating building data obtained from various height sources, which may affect all operations associated with estimating the number of floors.

8. Attribute Queries and Spatial Distribution Visualization
LOD1 building datasets of differing specifications may be linked to diverse attribute sources. Errors or inconsistencies in attribute accuracy can introduce risks in geographic spatial distribution analyses and visualization tasks.

In summary, without comprehensive usability metadata, users are unlikely to develop an accurate understanding of the acquired LOD1 building datasets, thereby compromising the appropriate application of the data within intended contexts. Even with advanced GIS software capabilities, the absence of intelligent mechanisms to assess data usability renders analytical outcomes susceptible to significant risks.

3.3 Standardized Metadata Encoding

This research employs GeoNetwork version 4.4.8 as a platform for the development and management of metadata, facilitating the creation, maintenance, and querying of geographic information metadata in accordance with the ISO 19115-3 standard. GeoNetwork is an open-source geographic information catalog system maintained by the OSGeo community. It can be seamlessly integrated with the also open-source GeoServer software to establish a more comprehensive network service infrastructure. The primary functionalities of GeoNetwork include metadata editing, validation, map visualization, and multi-protocol publishing, thereby addressing the diverse requirements of metadata creators, administrators, and end-users. Metadata management is conducted via a multilingual web interface that supports role-based access and permission controls. GeoNetwork provides a standardized interface for metadata editing, users may select the appropriate standard schema on the "Metadata & Templates" page to generate metadata content that adheres to selected international standards (e.g., ISO 19115-3), thereby underscoring the advantages of global standardization.

In this study, the standardized metadata constructed is exported in XML format, with its schema structure conforming to the schema of ISO 19115-3. Given its text-based digital encoding, the metadata content can be accessed and reviewed using any text editor or digital content management tool. Portions of the XML metadata generated in this research are presented in Figures 2. As previously noted, the overall metadata operation constitutes an ecosystem involving all stakeholders within the spatial data infrastructure. By selecting software that adheres to international standards, this study facilitates seamless integration with metadata content developed under the same standards. Consequently, the entire framework can be constructed upon shared standardized technologies that mutually reinforce one another, thereby enabling the practical implementation and further development of the research outcomes.


```
<mri:abstract>
  <gco:CharacterString>1. This dataset was produced by the National Land Surveying and Mapping Center of the Ministry of the Interior as part of the National Base Map initiative. It provides multi-dimensional building data constructed in accordance with the CityGML LOD1 specification. The data were primarily generated through photogrammetry or by integrating 2D maps with height information to provide realistic representations of buildings. 2. The definition of a building follows the specifications of the topographic data classification framework: a structure fixed to the ground with a roof, beams, and walls, used for commercial, industrial, or residential purposes. The data represent buildings in three-dimensional space. 3. The spatial units include buildings and building frames. The latter are based on the characteristics of the original source data and may not meet individual building requirements. Therefore, the dataset may not always use single buildings as the unit and does not include detailed semantic subdivision of indoor or structural components. 4. The exterior appearance of the 3D buildings currently provided is a simulated texture and does not reflect real-world appearances. 5. Building data are categorized into two types based on their source: photogrammetric data and map-derived data. -> Photogrammetric data are produced by using aerial photographs with appropriate overlap to construct a 3D representation. -> Map-derived data include 1:1000 topographic maps and the standard electronic map. The 1:1000 topographic map is based on aerial survey results and represents individual buildings in 2D according to ridge lines. The standard electronic map uses the concept of building blocks, where multiple buildings within an area may be represented as a single polygon due to generalization, and therefore cannot be treated as individual building units. -> Because of differences in data sources and accuracy, these three categories may contain variations in building representation. 6. The minimum mapping requirement for buildings in the 1:1000 topographic map and the standard electronic map is 2 m × 2 m (4 m²) or a minimum length or width of 5 meters. Buildings smaller than these thresholds will not appear in the dataset. 7. 3D considerations: LOD1 buildings are limited to horizontal and vertical planar representations and cannot accurately depict actual roof shapes. 8. This dataset does not include detailed objects such as doors or windows. 9. This dataset corresponds to the 2019 version.
  </gco:CharacterString>
</mri:abstract>
```


Figure 2 XML-based metadata - abstract information

3.4 Awareness Design at Users Software

This study presents the integration of metadata reading capabilities within the QGIS platform to enhance users' proficiency in understanding applied datasets. The proposed "automatic and awareness" methodology can be incorporated into operational workflow design, thereby promoting users' identification of essential domain-specific knowledge relevant to the data resources employed. QGIS, an open-source, cross-platform geographic information system, provides Python API (PyQGIS) that allows developers to programmatically manipulate various features, including map datasets, data processing, and spatial analysis. Its extensible plugin architecture enables the implementation of customized functionalities tailored to specific requirements, facilitating seamless interaction with existing datasets, tools, and user interfaces. The mechanism developed herein automatically parses XML-based metadata files compliant with the ISO 19115 standard, integrating critical information into the operation design workflow and thereby enhancing the capacity to assess data applicability through perceptual insights.

Certain operations are executed automatically, whereby users receive "automatic prompts" containing explanations regarding dataset usability. These explanations encompass various data design and construction parameters considered during earlier stages and, through comprehensive documentation, enable users to apprehend the status and application constraints of the acquired data. A representative example involves the immediate reading of fundamental domain information following data selection and loading, as such universe of discourse information influence all subsequent operational decisions related to the dataset. Upon selecting a LOD1 building dataset, the system automatically parsing the metadata, and salient abstract and lineage information is displayed within the user interface (Figure 3 and Figure 4).

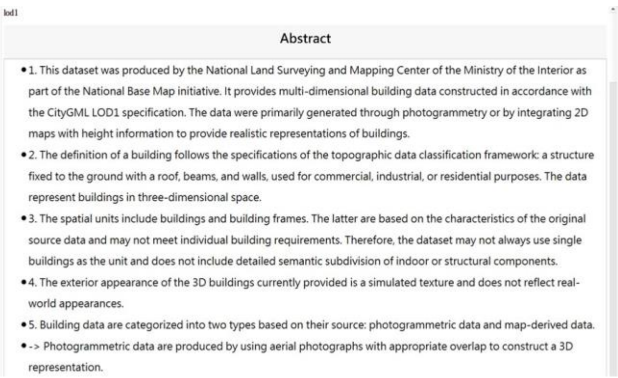

Figure 3 Awareness output – abstract information

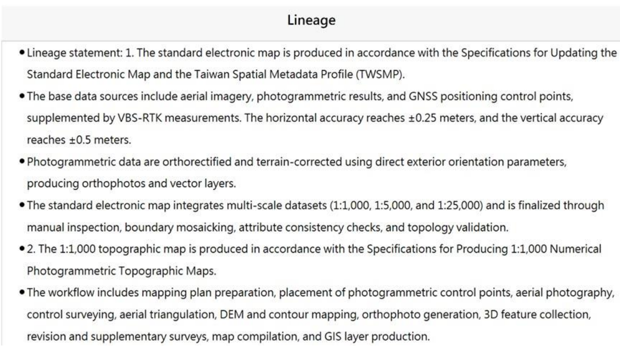

Figure 4 Awareness output – Lineage information

Interpretation functionalities are also embedded within workflow designs aligned with the operational characteristics of GIS software. For instance, when the AreaCalculation toolbar button is activated, the system first verifies the feature type. If the feature is not polygonal, an error message is presented and the calculation is aborted. Conversely, if the feature is polygonal, the dataset's usability-related information (UsabilityElement) is displayed in a pop-up window to promptly inform the user (Figure 5). The area calculation proceeds and the result are recorded in the area attribute only after users' confirmation.

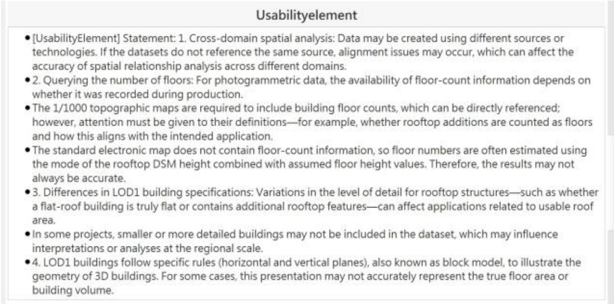

Figure 5 Awareness output-usability for Area calculation

4. Conclusion

In contrast to traditional geographic resource sharing methodologies that primarily emphasize resource discovery and access processes, this research seeks to develop an interdisciplinary sharing environment founded upon an "enhanced usability perspective." The central objective is to propose concrete strategies for articulating the "usability" characteristics of geospatial data through a "standardized framework." This framework facilitates data services and providers in delivering detailed descriptive content grounded in domain expertise and aligned with international metadata standards, thereby promoting efficient data sharing. By employing a standardized metadata-based mechanism, users are enabled to accurately comprehend the acquired resources, evaluate their "usability" within specific contexts and scenarios, and discern differences among various datasets to mitigate the risk of erroneous decision-making.

With respect to specific contributions, the metadata exploration advanced herein promotes the principles of interoperability and reusability. This is especially relevant in sharing environments encompassing extensive cross-domain resources, such as national spatial data infrastructures, where obtaining first hand descriptive information from service providers and applying it appropriately confers significant advantages to users. By enhancing the quality of geographic information sharing and utilization from a knowledge-centric standpoint, this study supports the advancement of future national base maps designed to underpin diverse domain applications. Further investigation into the usability of common features may facilitate the realization of multi-objective, cross-domain applications.

For software developers, the enhancement of professional competencies in data handling through improved functionalities and the integration of standardized metadata constitutes a critical pathway for innovation. In summary, the challenge addressed in this study represents a persistent issue within the geographic information domain. Without the development of concrete strategies, this problem is likely to exacerbate in tandem with the increasing diversity and volume of data. An effective professional division of labor, coupled with consensus on supporting standards, emerges as the most viable solution to this challenge.

https://doi.org/10.5194/isprs-archives-XLIX-B4-2026-49-2026 | © Author(s) 2026. CC BY 4.0 License.

Acknowledgements

Partial results of this study is sponsored by the research grant from National Science and Technology Council (NSTC 113-2121-M-006-018)

References

Aalders, H. J. 2005. An introduction to metadata for geographic information. In World Spatial Metadata Standards, pp. 3-27, Elsevier.

Ahonen-Rainio, P., & Kraak, M.-J. 2005. Deciding on fitness for use: evaluating the utility of sample maps as an element of geospatial metadata. *Cartography and Geographic Information Science*, 32(2), 101-112.

Barzaghi, S., Bordignon, A., Gualandi, B., Heibi, I., Massari, A., Moretti, A., … & Renda, G., 2024. A Proposal for a FAIR Management of 3D Data in Cultural Heritage: The Aldrovandi Digital Twin Case. *Data Intelligence*, 6(4), 1190-1221. https://doi.org/10.3724/2096-7004.di.2024.0061

Brodeur, J., Coetzee, S., Danko, D., Garcia, S. & Hjelmager, J. 2019. Geographic Information Metadata—An Outlook from the International Standardization Perspective. *ISPRS International Journal of Geo-Information*, 8（6）. doi:10.3390/ijgi8060280

Devillers, R. and Zargar, A., 2009. Towards quality-aware GIS: operation-based retrieval of spatial data quality information. ed. Spatial Knowledge and Information （SKI） Conference", Fernie, B.C., Canada, February.

Devillers, R., Y. Bedard, M. Gervais, Robert Jeansoulin, François Pinet, et al.. 2007a. How to improve geospatial data usability: from metadata to quality-aware GIS community. AGILE Pre-Conference working group on spatial data usability, May 2007, Aalborg, Denmark. pp.6. ⟨hal-02589086⟩

Devillers, R., Bédard, Y., Jeansoulin, R., & Moulin, B. 2007b. Towards spatial data quality information analysis tools for experts assessing the fitness for use of spatial data. *International Journal of Geographical Information Science*, 21(3), 261-282.

Devillers, R., Stein, A., Bédard, Y., Chrisman, N., Fisher, P., & Shi, W. 2010. Thirty Years of Research on Spatial Data Quality: Achievements, Failures, and Opportunities. *Transactions in GIS*, 14(4), 387-400. doi:10.1111/j.1467-9671.2010.01212.x

Fischer, J., Egli, L., Groth, J., Barrasso, C., Ehrmann, S., Figgemeier, H., Henzenc, C., Meyer, C., Müller-Pfefferkorn, R., Wagner, M., Bernard, L., & Seppelt, R. 2023. Approaches and tools for user-driven provenance and data quality information in spatial data infrastructures. *International Journal of Digital Earth*, 16(1), 1510-1529, DOI:10.1080/17538947.2023.2198778.

Giuliani, G, Cazeaux, H, Burgi, P-Y, Poussin, C, Richard, J-P and Chatenoux, B. 2021. SwissEnvEO: A FAIR National Environmental Data Repository for Earth Observation Open Science. *Data Science Journal*, 20: 22, pp. 1–18. DOI: *https://doi.org/10.5334/dsj-2021-022*

Hong, J-H & Huang, M-L 2017. Enabling smart data selection based on data completeness measures: a quality-aware approach, *International Journal of Geographical Information Science*, DOI: 10.1080/13658816.2017.1290251

ISO. 2014. ISO19115-1, Geographic information – Metadata - Part 1: Fundamentals. In: International Organization for Standardization.

ISO. 2023a. ISO19115-3, Geographic information – Metadata - Part 3: XML schema implementation for fundamental concepts. In: International Organization for Standardization.

ISO. 2023b. ISO 19157 Geographic information - Data quality Part 1 General requirement. In: International Organization for Standardization.

Ivánová, I., N. Brown, R. Fraser. N. Tengku and E.Rubinov, 2019. FAIR and standard access to spatial data as the means for achieving sustainable development goals. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-4/W20, 2019 ISPRS and GEO Workshop on Geospatially-enabled SDGs Monitoring for the 2030 Agenda, 19–20 November 2019, Changsha, China.

Manso-Callejo, M. Á., & Wachowicz, M. 2009. GIS Design: A review of current issues in interoperability. *Geography Compass*, *3*（3）, 1105–1124. https://doi.org/10.1111/j.1749-8198.2009.00241.x

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.I.J. et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018.